

Soluzione Esercizio 1

- Caricare entrambi i dataset in Google Refine - Aprire i due file in due tab diversi del browser
- Eliminare le colonne che non sono a comune, ad eccezione della colonna Tipologia Museo nel dataset della Lombardia.

nazione	Indirizzo	Ubicazione	Cap	Telefono	Fax	Email
L	Facet		1021	763733642	763731175	info@museodelfi
cittÀ	Text filter	zzo	1021	0763 7309205	0763 711215	museo.civ.dioc@
cittÀ	Edit cells	zzo	1023	0761 796914		info@simulabo.it
	Edit column					
	Transpose					
	Sort...					
	View					
	Reconcile					
REPUBBLICA; 29	Rev					
DE LA	Carr					
	Apo					
DI						
RA						
TESIO						
DE						
L						
ELLA						
LE E						
	PIAZZA	-	3012	775727856	775727856	-

- Aggiungere una colonna Tipologia nel dataset del Lazio
- Rinominare le colonne rimanenti in modo da avere lo stesso nome nei due dataset

ione	Indirizzo	Ubicazione	Cap	Telefono	Fax	Email
	Facet		1021	763733642	763731175	info@museode
À	Text filter	zzo	1021	0763	0763	museo.civ.dioc
À	Edit cells	ovile		7309205	711215	
	Edit column	zzo	1023	0761 796914		info@simulabo.
	Transpose					
	Sort...				09	comune_alatri@
	View					culturaeturismo
CO	Reconcile				59	museo@comur
	REPUBLICA; 29	Rev			93	museoallumiere
.A		Can				
	VIA COLA; 47	Apo			32	museocivico@c
IO						
	VIA LEONE XIII; 5	C/o			74	anagnicattedral
A		di A				

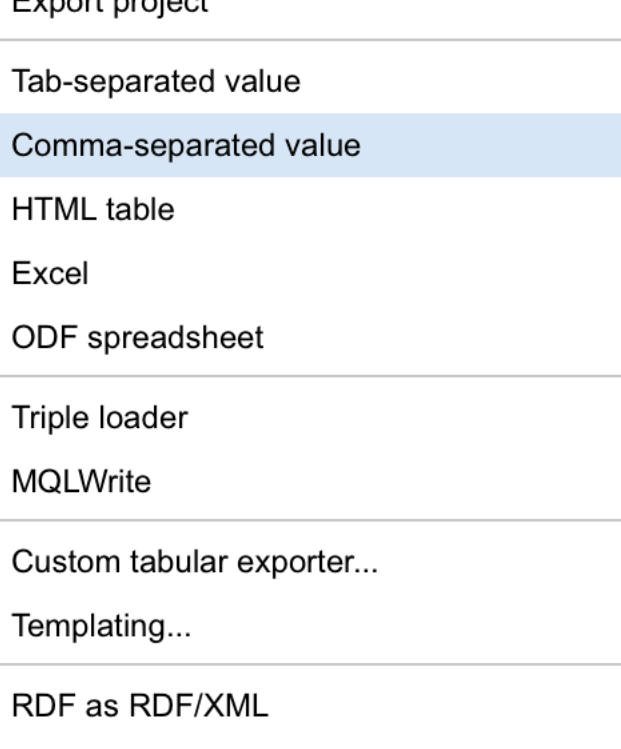
- Disporre il nome nelle colonne nello stesso ordine in entrambi i dataset. Per fare questo spostare le colonne

one	Indirizzo	Ubicazione	Cap	Telefono	Fax	Er
	Facet		1021	763733642	763731175	info@r
	Text filter	zzo	1021	0763 7309205	0763 711215	musec
	Edit cells	zzo	1023	0761 796914		info@s
	Edit column					
	Transpose					
	Sort...					
	View					
	Reconcile					
	REPUBLICA; 29	Rev Carr Apo				
	VIA COLA; 47	-				
	VIA LEONE XIII; 5	C/o di A				
	PIAZZA	-	3012	775727856	775727856	-

- Il campo SitoWeb del dataset della Lombardia contiene dei tag HTML che vanno rimossi. Per fare questo, selezionare add Column based on this column e digitare (vedi esercitazione Google Refine)

`substring(value,indexOf(value,">")+2,indexOf(value,""))`

- Eliminare la colonna originale del SitoWeb nel dataset della Lombardia
- Aggiungere una colonna Regione (add Column based on this column) e nel campo espressione inserire il nome della regione tra virgolette, ad esempio "Lombardia"
- Esportare i due dataset in formato CSV
- Accedere dal proprio browser a Phpmyadmin. Selezionare un database (ad esempio LPW) e scegliere importa. Selezionare il file da importare (ad esempio MuseiLombardia.csv) e mettere il segno di spunta alla voce "La prima linea del file contiene i nomi dei campi della tabella (se quest'opzione non é attivata, la prima linea diventerá parte dei dati)". Caricare il file
- Selezionare la tabella Table 1 sulla sinistra e selezionare nuovamente importa. Caricare il dataset MuseiLazio.csv
- Rinominare la tabella in Musei selezionando Operazioni, Rinomina.



Open... Export ▾ Help

Export project

Tab-separated value

Comma-separated value

HTML table

Excel

ODF spreadsheet

Triple loader

MQLWrite

Custom tabular exporter...

Templating...

RDF as RDF/XML

RDF as Turtle